DOCUMENT RESUME

ED 079 282 SP 006 692

TITLE	Evaluation Systems for Education: Descriptive Abstracts.
INSTITUTION	National Education Association, Washington, D.C. Div. of Instruction and Professional Development.
PUB DATE	Mar 73
NOTE	31p.
AVAILABLE FROM	Instruction and Professional Development, National Education Association, 1201 16th St., N. W., Washington, D. C. 20036
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Abstracts; Criteria; *Evaluation Criteria; *Performance Criteria; Teacher Associations; *Teacher Evaluation; *Teacher Rating; Teachers

ABSTRACT

This document contains brief abstracts of 13 different teacher evaluation systems; each includes a description of the system and its purpose, an outline of criteria, and an explanation of how it works and where to get further information. It is intended as an aid to teachers and teacher associations in assessing the appropriateness of a variety of evaluation systems. (JB)

ED 079282

EVALUATION SYSTEMS FOR EDUCATION:

DESCRIPTIVE ABSTRACTS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SP 006 692

Program and Performance Evaluation Team

INSTRUCTION AND PROFESSIONAL DEVELOPMENT
NATIONAL EDUCATION ASSOCIATION
1201 - 16th Street, N. W., Washington, D. C. 20036

March 1973

FILMED FROM BEST AVAILABLE COPY

This document is not copyrighted and may be duplicated without permission, provided credit is given to the National Education Association.

CONTENTS

Preface 5

Educator Feedback 7

Flanders System
of Interaction Analysis 9

Indicators of Quality 11

Instrument for the
Observation of Teaching Activities. 13

OScAR 5V. 15

Performance Assessment Record
for Teachers. 17

Profiles of Excellence. 19

Purdue Rating Scale
for Administrators and Executives 21

Quality Inventory 23

Reciprocal Category System. 25

Teacher Performance Appraisal 27

Teaching Performance Tests. 29

33 Roles for Teachers and Pupils
in the Classroom. 31

Members of the IPD Program and Performance Evaluation Team are Bernard H. McKenna, Secretary; Richard W. Cortright, Roy A. Edelfelt, Thelma Horacek, Margaret Knispel, Geraldine E. Pershing, and Frances Quinto.

PREFACE

Probably every teacher in the United States will be evaluated. Already, scores of teachers have been or are being evaluated, sometimes fairly and adequately, sometimes wrongly and perversely. Teachers are bringing the problems of performance evaluation to their local education association officers, to UniServ directors, and to their state associations and the National Education Association.

To give support to teachers in the arena of staff performance and school program evaluation, the NEA, through its office of Instruction and Professional Development, has established a Program and Performance Evaluation Team. The team has prepared this booklet to help teachers and teacher associations assess the appropriateness of a variety of evaluation systems. It contains brief abstracts of thirteen different systems, each including a description and purpose of the system, an outline of criteria, and an explanation of how it works and where to get further information.

The publication of this document does not represent endorsement by the NEA of any of the systems described. It does suggest, however, that each of the systems has been researched and developed over a considerable period of time and deserves consideration by teacher associations as they become involved in making determinations about how evaluation will take place.

Original compilation of the abstracts was made by Jeanne Wine. The copy was prepared by Karen Bonner.

Richard W. Cortright
Bernard H. McKenna

EDUCATOR FEEDBACK (TEACHER IMAGE QUESTIONNAIRE)
(William D. Coats)

DESCRIPTION AND PURPOSE: The major objective of Educator Feedback is to provide teachers with confidential information designed to help them work more effectively with students.

EVALUATION CRITERIA: Each student in a class is provided with an 18-item questionnaire to measure his perception of the teacher's effectiveness, responding to each item on a five-point scale ranging from poor to excellent. For example:

1. Knowledge of subject: Does he have a thorough knowledge and understanding of his teaching field?
2. Fairness: Is he fair and impartial in his treatment of students?
3. Variety of teaching procedures: Is much the same procedure used every day, or are different and appropriate teaching methods used at different times?
4. Sense of humor: Does he see and share with students amusing happenings and experiences?

INTERPRETATION AND APPLICATION OF RESULTS: The Educator Feedback Center summarizes responses to the questionnaire for the teacher, who receives a Teacher Image Profile, an interpretive discussion of factors which might be causing problems identified in his Profile, and suggestions for possible behavioral changes that might improve his effectiveness. With this system the teacher can also determine (a) his progress, by comparing several profiles across time, (b) his effectiveness with several groups through profiles from different classes, and (c) discrepancies among his self-image, ideal image, and real image with respect to teaching effectiveness.

HOW IT WORKS: The Teacher Image Questionnaire takes 10-15 minutes to complete. No training in classroom observation is required for this system. The cost of processing questionnaires and preparing a teacher profile for one class of twenty-five is $7.

FOR FURTHER INFORMATION: Educator Feedback Center, School of Education, Western Michigan University, Kalamazoo, Michigan 49001.

FLANDERS SYSTEM OF INTERACTION ANALYSIS (FSIA)
(Ned Flanders)

DESCRIPTION AND PURPOSE: The Flanders System of Interaction Analysis is an observation system that focuses on teacher talk and student talk in the classroom. (A revised form of FSIA, called the Flanders Expanded System, is also available.) FSIA has been used in a variety of teacher training programs to provide trainees with personal feedback about teaching behaviors and their effect on the quantity of student participation in the classroom.

EVALUATION CRITERIA: There are seven categories of teacher talk divided into "indirect" and "direct" influence on students -- accepts feelings, praises or encourages, accepts or uses ideas of student, asks questions, lectures, gives directions, and criticizes or justifies authority. And there are three student categories -- response initiation and silence or confusion. Results of the use of FSIA indicate high observer reliability.

The FSIA allows the evaluator to tell by looking at the matrix of categories what preceded and what followed every verbal behavior of both the teacher and the students.

HOW IT WORKS: The FSIA can be used in either elementary or secondary school and for any subject matter. An evaluator can administer the system by himself. No special equipment is needed to code behavior. Observations may be live or from video or audio tapes. Minimum study of the system is needed before using it.

FOR FURTHER INFORMATION: Addison-Wesley Publishing Company, Reading, Massachusetts 01867.

INDICATORS OF QUALITY

DESCRIPTION AND PURPOSE: Indicators of Quality is an instrument for obtaining quantitative measures of school quality by observation of critical behavior in the classroom. Its primary purpose is to serve as a criterion in school quality studies.

EVALUATION CRITERIA: Indicators of Quality is based on four characteristics of internal school behavior that are judged by several years of research and field testing to be basic quality criteria: individualization, interpersonal regard, creativity, and group activity.

The instrument consists of fifty-one items derived from the four basic criteria, each with a positive and negative extreme of classroom behavior or activity (polar characteristic). Of these items, seventeen are observable in teacher behavior, seventeen in pupil behavior, and seventeen in teacher-pupil interaction. Observers obtain a series of time samples of standard length and structure, following precise instructions in timing attention to teacher, pupil, and teacher-pupil interaction. Schedules are set up to obtain a sampling of all class meetings in the school throughout the observation day. These samplings are stratified by subject matter (in secondary school), time of day, pupil ability, size of class, sex of teacher, and teacher in charge (whether regular, substitute, student, or specialist teacher).

INTERPRETATION AND APPLICATION OF RESULTS: All data obtained from Indicators of Quality are analyzed and interpreted by the authors for a school district. At the local school district level, results of the application of Indicators

of Quality can serve in the diagnosis of areas for school improvement. As the instrument is applied to groups of school systems throughout the country, the resulting normative data provide a frame of reference against which a given system can measure itself. As more is learned about the organizational conditions under which high scores are obtained (size of class, grade level, subject, style of teaching, etc.), information can be obtained about how local conditions can be effectively modified. The authors emphasize that the use of this system results in a total school district score and that its use in discriminating among individual teachers has not proved promising.

HOW IT WORKS: Teaching is administered in a carefully structured, three-day period of familiarization and trial application. During the first six days of an observer's work, a reliability check with other observers is conducted in order to identify highly variable or grossly divergent observers. In addition, observation schedules are constructed by a special coordinator since consistency in the sampling of class meetings is a requisite.

Indicators of Quality materials are not available for general distribution. A complete application, scoring, and report service entailing the use of outside trained observers is available. Observer training requires three days, coordinator training, one day. Total training and 350 observations cost $3500, plus travel for one trainer. Trainers, are available in most regions of the country.

FOR FURTHER INFORMATION: Vincent and Olson School Evaluation Services, Box 22, Salisbury, Connecticut 06068.

INSTRUMENT FOR THE OBSERVATION OF TEACHING ACTIVITIES (IOTA)

DESCRIPTION AND PURPOSE: The IOTA is an instrument for evaluating teacher competence against a standard of excellence as defined by the National IOTA Council. (Council members are educators.) Classroom observations and structured interviews are utilized to collect information in order to construct a profile of a teacher. The teacher's performance is measured against accepted criteria rather than against other teachers.

EVALUATION CRITERIA: There are twenty-seven scales, each representing one factor considered essential to effective teaching and learning. Fourteen observation scales measure such areas as variety in activities, classroom control, and individualization of instruction. Thirteen interview scales measure such areas as participation in school staff activities, utilization of community resources, and skill in enhancing multicultural relationships.

INTERPRETATION AND APPLICATION OF RESULTS: IOTA is to be used only for the improvement of instruction through:

1. Teacher orientation to professional growth, providing a common frame of reference of teaching excellence.
2. The development of a common point of view between administration and teachers as to the philosophical and operational objectives of teaching.
3. The encouragement of teacher self-evaluation and self-analysis leading to self-improvement.
4. The provision of more objective and accurate appraisal of teaching competence by administrators which is based on a commonly accepted frame of reference.
5. The provision of a communication vehicle for open discussion among teachers and between teachers and administrators.

HOW IT WORKS: School districts wishing to use the IOTA system must be approved by the National IOTA Council, and qualified directors and consultants must be available to the districts. The instrument must be used only for the improvement of instruction. (For example, the Council will not approve use of the instrument for salary determinations.) It can be used to evaluate teachers at any grade level or in any subject area. The program of training must meet the current specifications of the National IOTA Council. Only those authorized by the Council may train observers

FOR FURTHER INFORMATION: Dr. R. Mertwin Deever, Director, Bureau of Educational Research and Services, College of Education, Arizona State University, Tempe, Arizona 85281.

OScAR 5V (OBSERVATION SCHEDULE AND REPORTING FORM, NO. 5, VERBAL)

(Donald M. Medley)

DESCRIPTION AND PURPOSE: OScAR 5V is an instrument for classifying and recording a teacher's verbal behavior in the classroom. During the period of observation, every verbalization the teacher makes is recorded, along with any pupil verbalizations to which the teacher responds directly. Teacher verbalizations are classified both by frequency and by affective, cognitive, or procedural content; pupil verbalizations are recorded in much less detail.

EVALUATION CRITERIA: There are fourteen types of verbalizations used in coding teacher verbal behavior in the OScAR 5V, such as the nonsubstantive question, rebuking statement, and convergent question. Pupil verbalizations are of four types, such as pupil question and pupil statement. Simple objective criteria are employed for discriminating categories from each other. However, the coder must at times classify ambiguous teacher statements on the basis of how they are received by the student rather than on the teacher's intent. As a result, the discrepancies between records made by different observers may be attributed to ambiguities in the behaviors themselves rather than to observer errors.

INTERPRETATION AND APPLICATION OF RESULTS: The completed OScAR 5V answer sheets can be interpreted to some extent by visual inspection. The record sheets provide a detailed sequential account of the teacher's performance. Another method of interpretation is an analysis of the frequency of occurrence of categories during a particular sequence. These two methods can be used by supervisors to:

1. Demonstrate objectively to a teacher the presence or absence of desirable or undesirable teaching strategies.

2. Evaluate a teacher's progress both before and after attempts to change behavior have been instituted.

3. Relate teacher behaviors to pupil learning or some other criterion of success in teaching.

4. Define immediate goals for a teacher in training.

5. Assess the effects of a curriculum change on a teacher's classroom behavior.

HOW IT WORKS: Neither the length nor the frequency of observation is specified. Little accuracy is lost if, for convenience, one codes from audio tapes or typewritten transcripts from the classroom, although live observation is preferable.

The mechanics of coding OScAR 5V are specified in detail in the manual. However, since the sixty-eight possible combinations of classifiable behavior, as well as the system itself, may be confusing to the untrained observer, a minimum of two but preferably four days with a trainer is recommended. Teachers or other trained personnel familiar with the classroom environment and able to quickly analyze complex behavior into simple, objective components would make competent observers.

The three items of expense to school systems using OScAR 5V are observer training, record sheets, and machine scoring. To facilitate interpretation, it is recommended that behaviors be recorded on a machine-scorable answer sheet from which the record may be transferred to punched cards or magnetic tape and then fed into a computer to count frequencies. However, hand scoring is possible.

FOR FUTHER INFORMATION: Donald M. Medley, Bureau of Educational Research, Peabody Hall, University of Virginia, Charlottesville, Virginia 22903.

PERFORMANCE ASSESSMENT RECORD FOR TEACHERS (PART)
(Bob Burton Brown)

DESCRIPTION AND PURPOSE: PART is an instrument for systematically observing, describing, and assessing the classroom performance of teachers and pupils. The observer watches and listens for the twenty-two behaviors described in the instrument and records them as they occur during each of five separate six-minute recording periods. PART can also be used by teachers in self-assessment.

EVALUATION CRITERIA: The first four items on PART measure the extent to which classroom activity is inquiry-centered, or the extent to which students are involved in reflective thinking. Items 5-15 measure the level of cognitive or thinking operations used overtly by students in the classroom. Items 16-19 are concerned with the students' value systems, including both the basis for and the process involved in conceptualizing, organizing, and intellectualizing values. The last three items represent teacher control factors that have been found to relate most strongly to student achievement and are based on factor analysis of a variety of observational systems. One such item is related to freedom without the extremes of too much or too little control.

INTERPRETATION AND APPLICATION OF RESULTS: The items represent a cross-section of ways of analyzing teaching-learning situations in classrooms. The instrument is designed specifically to serve as a framework for looking at and thinking about teaching in more than customary depth for the sole purpose of improving its quality. PART does not lend itself well to research designs or to teacher evaluation schemes which ascribe a single qualitative or quantitative description to teaching performance.

HOW IT WORKS: Ten hours of training in five two-hour sessions are recommended in order for observers to achieve an adequate level of reliability. Teachers can be effective observers of video tapes of their own behavior, or observations may be carried out in the classroom by trained observers. Training costs for groups of up to fifty persons range from $1500 to $2500, depending on distance traveled and time required by the trainers.

FOR FURTHER INFORMATION: Institute for Development of Human Resources, College of Education, University of Florida, Gainesville, Florida 32601.

PROFILES OF EXCELLENCE
(National Education Association)

DESCRIPTION AND PURPOSE: Profiles is a program evaluation instrument designed for the study and evaluation of a variety of elements in the total operation of a school system. Its primary purpose is to assist local professional associations, boards of education, and administrators in identifying areas for improvement. It is not be be used as a measure of the quality of teaching, an accrediting instrument, or a means of establishing comparative standards among systems. The focus of Profiles extends across the affective, cognitive, procedural, structural, and environmental dimensions.

EVALUATION CRITERIA: The criteria for evaluating a school system are based on scientific evidence, professional judgment, and exemplary practice. The instrument consists of nine main sections:

The Educational Program
Administrative Operation
Board of Education
Staff Personnel Policies and Procedures
Professional Compensation
Conditions of Professional Service
The School Plant
District Organization, Finance and Business Administration
The Local Professional Organization.

Each section has subtopics and each subtopic consists of two descriptive paragraphs, one labeled "inferior," the other "ideal." A school system is rated on a continuum between these extremes. Included as a part of each subtopic is the requirement for the evaluator to list specific recommendations for action to move the school system's rating to the superior category in a particular operational aspect.

INTERPRETATION AND APPLICATION OF RESULTS: An association using Profiles should appoint a coordinating committee to oversee several subcommittees which will independently evaluate an allotted number of subtopics. Each subtopic should receive a minimum of three independent evaluations. The various subcommittee reports are to be submitted to the coordinating committee for incorporation into a final evaluation report which may then be used as an authoritative basis for efforts to achieve needed improvements.

HOW IT WORKS: A full-scale school system evaluation will take six months. Members of the coordinating committee should be chosen for their depth and breadth of educational experience, and subcommittee members need to be dependable and thorough. The Profiles handbook costs $1.25. All other expenses will be generated within the school system.

HOW TO ORDER: Order Profiles of Excellence (stock No. 381-11774; $1.25) from Customer Service Section, National Education Association, 1201 16th Street, N. W., Washington, D. C. 20036.

PURDUE RATING SCALE FOR ADMINISTRATORS AND EXECUTIVES
(ELEMENTARY AND SECONDARY SCHOOL ADMINISTRATORS FORM C)
(H. H. Remmers and R. L. Hobson)

DESCRIPTION AND PURPOSE: The Purdue Rating Scale provides an administrator with a means for measuring the strengths or weaknesses of his own performance through the anonymous ratings given to him by those for whom he is responsible. It can be used in the public schools by any person in a supervisory capacity, such as superintendent, principal, department head, supervisor, or head teacher. The instrument is recommended solely as a basis for self-improvement.

EVALUATION CRITERIA: The Purdue Rating Scale is a list of thirty-six statements about administrators sub-grouped into ten categories:

Intellectual Balance
Emotional Balance
Administrative Leadership
Administrative Planning
Use of Funds
Capacity for Work
Accomplishment
Relations with Subordinates
Public Relations
Social Responsibility.

INTERPRETATION AND APPLICATION OF RESULTS: Each statement is scored separately and its mean plotted on a Profile Chart. There are three different kinds of interpretation of the Profile Chart:

1. Literal interpretation, where the average score of each item is studied to find how an administrator's subordinates as a group view him.

2. Comparison of the administrator's ratings with a national, normative group of public school superintendents and principals. The percentiles yielded by these normative ratings are printed on the Profile Chart for ease in comparing a particular score to the norm.

3. Distributive interpretation, which reveals how many employees gave the administrator each of five possible ratings on each item, and how much agreement there was among the ratings.

Scoring the scales and reporting the results can be done confidentially by the Personnel Evaluation Research Service at Purdue University.

HOW IT WORKS: It is recommended that use of the Purdue Rating Scale be voluntary and that all results be confidential. The scale is self-administering and requires only a few minutes to complete.

HOW TO ORDER: Order by the package -- 25 rating sheets for $3.50 (discounts on quantity orders) -- or a specimen set for $1.00, from University Book Store, 360 State Street, West Lafayette, Indiana 47908.

QUALITY INVENTORY
(W. S. Vincent)

DESCRIPTION AND PURPOSE: Quality Inventory is a rating-scale type instrument for appraising a variety of aspects of a school district program. Individual schools and the school system are rated against two reference groups of school systems. Comparative data are based on recent educational research or expert judgments. The Quality Inventory does not automatically generate a plan for improving a school system, although it does suggest areas where improvement should be considered as related to ranges and norms for the reference systems on the various criteria.

The Quality Inventory may be used alone but it is also part of a package that includes the instruments Indicators of Quality (pp. 11-12) and Index of Pupil Gain. The former obtains scores from samplings of observations of educational activities and the latter samples output by means of achievement tests results. The Quality Inventory was designed to help a school system discover ways to improve scores on both of these instruments.

EVALUATION CRITERIA: The items on the Quality Inventory are organized into categories of program conditions: finance, professional staff, curriculum, plant and facilities, and community relations.

INTERPRETATION AND APPLICATION OF RESULTS: The results from applying the various criteria are either quantifiable (e.g., number of staff per 1,000 students) or based on expert judgment. In a number of instances means and ranges are provided for a national group of school systems and for a suburban group, making possible a comparison of findings. Suggestions are made for action based on the results.

HOW IT WORKS: It is recommended that the Quality Inventory be applied by a team of administrators, teachers, or a combination of staff. To obtain data on some criteria, such as equipment or educational practices, direct inspection or interviews may be necessary. The results can be combined into a comprehensive analysis of the entire school district. Training is not required to apply the instrument, but the assistance of school business office personnel may be required in identifying data and making computations. Consultation on analysis, interpretation of results, and follow-up action is available, beginning at $200 a day. The Quality Inventory costs $6 per copy, with discounts on quantity orders.

FOR FURTHER INFORMATION: Vincent and Olson School Evaluation Services, Box 22, Salisbury, Connecticut 06068.

RECIPROCAL CATEGORY SYSTEM (RCS)
(Richard L. Ober)

DESCRIPTION AND PURPOSE: RCS is an observation system designed for teachers, supervisors, and administrators to obtain objective feedback concerning verbal behaviors in classroom teaching-learning situations. The observer keeps a running tally of behaviors that occur during successive three-second intervals. The system measures affective behaviors by producing descriptive data on the socioemotional climate, for example, teacher-student and student-student interaction, rigid-open social climate, supportive-restrictive teacher and student behavior. In addition, it produces data that specify certain cognitive levels of behavior, such as general versus specific questions.

EVALUATION CRITERIA: A modification of the Flanders System of Interaction Analysis (p. 9), the RCS consists of nine verbal behavior categories, each of which can be assigned to either teacher or pupil talk, plus one general category. When a verbal behavior is observed as teacher talk, its category is recorded as a single-digit number. The corresponding reciprocal number is used to record the same behavior by a student. The categories are:

1 or 11 "warms" (informalizes) the climate
2 or 12 accepts
3 or 13 amplifies the contributions of another
4 or 14 elicits
5 or 15 responds
6 or 16 initiates
7 or 17 directs
8 or 18 corrects
9 or 19 "cools" (formalizes) the climate
10 silence or confusion.

INTERPRETATION AND APPLICATION OF RESULTS: The observer, who in most cases is the teacher, learns to interpret the collected data so that he can view his own

teaching in an objective fashion. The author suggests that the system be used as a means of securing feedback so that the teacher can carry on his own self-evaluation and self-improvement program. In addition, RCS is useful as a tool both for collecting research data and for training teachers. When used as a research tool, the introduction of the reciprocity factor stimulates the conceptualization and measurement of more different verbal behavior patterns than might be possible with other instruments. Similarly, as a training tool, the reciprocity factor encourages the identification of a great number of verbal strategies that might otherwise escape the imagination and attention of the trainee.

HOW IT WORKS: Data may be collected by several modes: the observer may record (audio or video) his own teaching behavior and then collect data on playback, he may have a trained observer visit his classroom and collect data which he can study later, or he may use a combination of these techniques. RCS is applicable in any situation or subject area where there is verbal interaction.

Normally, a teacher-observer can learn to use the system with from 10 to 15 hours of training, depending on the level of sophistication required. If a workshop mode of training is followed, two or three days are required to prepare a group of observers. Normally a trainer charges a consultant fee ($100-$150 per day), plus expenses. Materials for training are provided by the trainer, with the exception of textbooks that might be needed.

FOR FURTHER INFORMATION: Richard L. Ober, College of Education, University of South Florida, Tampa, Florida 33620.

TEACHER PERFORMANCE APPRAISAL ("THE DENVER PLAN")
(Denver Public Schools)

DESCRIPTION AND PURPOSE: The purpose of the Teacher Performance Appraisal system is to encourage effective behavior by teachers. It assumes (a) that the key to excellence in education is the effectiveness of the teacher, and (b) that the teacher should continually assess the effectiveness of his teaching against educationally defensible, predetermined goals and then adjust his performance as needed to assure excellence of results.

EVALUATION CRITERIA: The teacher and the appraiser accept equal responsibility for carrying out the appraisal process, which involves the following steps:

1. Stating mutually agreed upon and measurable goals
2. Devising means of accomplishing the goals
3. Determining methods of evaluating progress toward those goals
4. Planning for communication between the teacher and the appraiser
5. Keeping permanent records
6. Summarizing progress in appraisal statements by the principal.

Goal-setting initially is based on the appraiser's knowledge of the needs of new teachers and then moves from the appraiser to the teacher during the appraisal period. Goals center around classroom management and planning, use of time and materials, and other areas in which new teachers need to develop skills rapidly. Subsequent appraisals reflect input from the teacher. At the conclusion of the appraisal period, the teacher should be an independent, goal-setting professional.

INTERPRETATION AND APPLICATION OF RESULTS: This method of teacher appraisal was designed to be a systematic tool (for measuring acceptable teacher performance) to replace the more subjective means that school principals may

currently use. The end result is a summary of progress, with comments from both the teacher and the appraiser and a final recommendation from the principal based on the teacher's ability to set meaningful goals, devise means of accomplishing the goals, and measure progress.

HOW IT WORKS: The principal requires assistance from administrative personnel assigned to his school, curriculum specialists, other principals, supervisors from the central administrative office, or from tenure teachers. The latter will usually be department or grade-level chairmen assigned to his school. The final responsibility for appraisal, however, will remain with the principal. There are no appraisal forms or answer sheets recommended for this method of evaluation and no fees involved.

FOR FURTHER INFORMATION: Harold A. Stetzler, Assistant Superintendent, Denver Public Schools, 414 Fourteenth Street, Denver, Colorado 80202.

TEACHING PERFORMANCE TESTS
(James Popham)

DESCRIPTION AND PURPOSE: Teaching Performance Tests were designed to measure teacher effectiveness without the flaws inherent in administrator ratings, pupil ratings, observation systems, and student performance on standardized tests. According to the author, traditional ratings are often based on highly variable conceptions of what constitutes good teaching and do not correlate well with most measures of pupil achievement. Moreover, student performance on standardized tests discriminates among learners rather than teachers and does not necessarily indicate the degree to which identifiable skills have been mastered.

Teacher Performance Tests avoid the measurement problems that arise when different teachers pursue different objectives. All teachers involved are asked to achieve the same objectives but are permitted to employ their own pedagogical techniques, making it possible to contrast teachers with respect to their skill in accomplishing identical goals.

EVALUATION CRITERIA: In the Teaching Performance Tests the teacher is given one or more explicit instructional objectives (and, preferably, a sample of the measurement procedure used to assess each objective), plus necessary related background information. He is provided sufficient time to plan an instructional sequence to accomplish the objectives and then instructs a group of learners. At the conclusion of the instruction the learners are measured on their achievement of the objectives, their performance providing an estimate of the teacher's instructional skill.

INTERPRETATION AND APPLICATION OF RESULTS: Teaching Performance Tests can be used for both instructional improvement and teaching skill assessment. As a tool for instructional improvement, they might be used by a school building unit, school system, or professional association in teaching performance clinics. For example, a team of teachers might meet to observe one teacher carry out a previously planned lesson with randomly selected learners. After the learners have been tested on their attainment of objectives and their affective responses to the teaching, the team can clinically analyze the teacher's instructional decision making and share demonstrably effective teaching tactics among colleagues.

The Teaching Performance Tests can also be used by professional associations, administrators, and schools of education for assessing a teacher's skill against established criteria.

HOW IT WORKS: At present, the author supplies those interested in Teaching Performance Tests with no more than the idea and rationale of their use. However, standardized packages are becoming available through other agencies Instructional Appraisal Services offers school districts a complete service for the screening of job applicants via the Teaching Performance Tests. The district supplies only the physical facilities for the testing, plus a limited number of students for a series of short enrichment lessons; the main cost of the testing is borne by the job applicant. The IAS supplies complete supervision of the testing program, plus, at its conclusion, a ranking of teacher applicants according to their skill in promoting different types of cognitive objectives as well as positive learner effect.

FOR FURTHER INFORMATION: W. James Popham, Instructional Objectives Exchange, University of California, Los Angeles, California 90024.

33 ROLES FOR TEACHERS AND PUPILS IN THE CLASSROOM

DESCRIPTION AND PURPOSE: 33 Roles is not an observation system in the traditional sense. While classroom observations are made, they are recommended as the basis for discussion among participants in the Role Identification Training Session. The observation instrument is designed to identify the joint roles that pupils and teachers play in teaching and learning. One purpose of 33 Roles is to increase a teacher's awareness of effective classroom behavior by observing examples of good teaching and learning. The 33 Roles are based on four principal objectives of education: individualization of instruction, good interpersonal relations among pupils and between teachers and pupils, cultivation of a setting for creativity and divergent thinking, and utilization of the characteristics of good group activity. During actual classroom sessions, observers using this instrument look for behavioral examples of pupils and teachers engaged in each of the thirty-three roles defined.

EVALUATION CRITERIA: Typical of the thirty-three roles described are:

1. Using a variety of resources and tasks
2. Mutual respect
3. Creative thinking
4. Question-and-answer technique
5. Individual communication.

INTERPRETATION AND APPLICATION OF RESULTS: The results of 33 Roles depend on the effectiveness of the Role Identification Training Session of which the instrument is only a part. There is no formal scoring system associated with the instrument, and its use is limited to those sessions.

HOW IT WORKS: Observers may be either teachers or pupils. 33 Roles could be used by teachers at any grade level or in any subject area. However, pupils younger than seventh grade age might not have the cognitive skills to behaviorally define the roles.

Three days of training are required. The total training cost for seventy-two participants, including all materials, is $1500, plus travel for trainers (two required). Trainers are available in most regions of the country.

FOR FURTHER INFORMATION: Vincent and Olson School Evaluation Services, Box 22, Salisbury, Connecticut 06068.